On Reporting Attitudes: an Analysis of Desire Reports and their Reading-Establishing Scenarios

Amsterdam Colloquium 2015
Tillmann Pross, prosstn@ims.uni-stuttgart.de
University of Stuttgart

December 2015

Summary: I argue that anaphoric constructions but not scope and intensional status of noun phrases are decisive for the meaning of attitude reports.

1 Three concerns about the conventional wisdom analysis of attitude reports

1.1 The conventional wisdom analysis

(1)   (A reporter says:) Adrian wants to buy a jacket like Malte's.

(A) The meaning of an attitude report like (1) finds expression in the readings it has

(B) The existence of readings is established by the difference that scenarios make to the interpretation of an attitude report: "[...] construct scenarios that make one of the readings true and the other false. This establishes the existence of two readings." (von Fintel and Heim, 2011, p. 85)

- Call such scenarios that entail the truth of an attitude report when it is taken in one way, but not when it is taken in another *reading-identifying scenarios*

- For example, a scenario in which Adrian has decided to buy a certain jacket but has no idea that the jacket he wants to buy is like Malte's gives rise to the so-called *de re* reading of (1): there is a jacket like Malte's which Adrian wants to buy.

- Or, in a scenario in which Adrian has not decided which jacket he wants to buy but he wants it to be like Malte's, (1) receives a *de dicto* interpretation, according to which Adrian wants to buy something that is a jacket like Malte's.

(C) The distinction between the *de re* and *de dicto* reading of an attitude report like (1) can be explicated in its logical form in terms of the scope relation between the modal verb *want* and the existentially quantified noun phrase *a jacket like Malte's*, as shown in (2).

(2)   a.   De re: $(\exists x)$(jacket(x) & like-Malte's-jacket(x) & wants(Adrian, buy(Adrian,x)))
       b.   De dicto: wants(Adrian, $(\exists x)$(jacket(x) & like-Malte's-jacket(x) & buy(Adrian,x)))

1.2 Three concerns about conventional wisdom

1.2.1 Fodor's puzzle

- Fodor (1970) argued that (B) predicts more readings of (1) than (C) can account for:

- In a scenario in which Adrian has decided what kind of jacket he wants to buy but has no idea that the kind of jacket he wants is like Malte's jacket, (1) intuitively has a reading which neither the *de re* nor the *de dicto* reading in (2-a), (2-b) capture correctly.

- The existential quantifier introduced by *a jacket like Malte's* must be inside the scope of *wants*, since there is no one particular jacket that Adrian wants to buy.

- But that would bring the description *a jacket like Malte's* within the scope of *wants* and that seems wrong since Adrian is assumed to know nothing about Malte's jacket.

- But the noun phrase *a jacket like Malte's* cannot be both inside and outside the scope of the verb *want*, thus Fodor's reading constitutes a paradox of scope.

1.2.2 First-Person Reports

- Given (A) and (C), why does a first-person report as in (3) leave no room for the traditional ambiguities?

- (3) cannot be a true utterance by Adrian if Adrian isn't fully aware that what he wants to buy is a jacket like Malte's.

(3) (Adrian says:) I want to buy a jacket like Malte's.

- The only difference between (3) and (1) is the grammatical subject of the report, thus conventional wisdom about the role of scope and intensional status of the indefinite noun phrase *a jacket like Malte's*, i.e. (C), cannot explain why (3) doesn't have the same variety of readings as (1).

- But if scope and intensional status of a noun phrase cannot account for why (1) allows for several readings while (3) only has one, then these are apparently not the right tools for dealing with these sentences.

1.2.3 Interpretation of Attitude Reports

- Given (A) and (B), how much must an interpreter of (1) know about the scenario which forces a certain reading upon it?

- An utterance of (1) would be uninformative for the interpreter in any situation where he understands enough about the scenario to be able to disambiguate (1) to its intended reading because according to (B) every reading-identifying scenario of (1) entails the truth of (1).

1.3 My diagnosis

- These three problems with the conventional wisdom about the meaning of attitude reports are all symptoms, I want to argue, of an underlying mistake: that of lumping together theorizing about the attitudes themselves with theorizing about the meaning of attitude reports.

2 Intra- and interpersonal anaphora in attitude reports

2.1 Intrapersonal anaphoric constructions

- Maier (2015), building on an idea in Kamp (1985) argues that non-doxastic attitudes such as desires are *parasitic* – i.e. 'anaphorically' dependent – on their doxastic host attitudes.

- For example, if Adrian believes that the object in front of him is a jacket and decides that he wants to buy *it*, the object of his want is anaphorically dependent on the object he believes to be a jacket.

- Attitudes of one and the same attitude holder that are *internally* about the same object (adopting the terminology of Kamp et al. (2011)) are connected by *intrapersonal* anaphoric links.

2.2 Interpersonal anaphoric constructions

- Attitudes can also be *externally* about the same object when "two expressions refer to the same real world entity, but where a particular speaker (or thinker) may be unaware of this" (Kamp et al., 2011, p. 332).

- Consider an example scenario for Fodor's puzzling reading in (4) (italics added and naming adopted to the running example of the paper).

(4) Reading-identifying scenario for (1) (Romoli and Sudo, 2009, p. 427)
Suppose a store sells some jackets that all look like Malte's and that Adrian does not know anything about Malte. Assume further that Adrian wants one of *those* jackets and any of *them* is an option.

- In (4), the plural pronoun *those* makes explicit that Adrian's attitude is externally about the same object of reference that is identified by the description *a jacket like Malte's*.

- Attitudes that are externally about the same object are connected by *interpersonal* anaphoric links.

- The crux of the use of such interpersonal anaphoric constructions in reading-establishing scenarios is the presentation of the same object of reference from two distinct points of view.

- (Other scenarios that establish Fodor's reading employ different anaphoric constructions, e.g. coreferring proper names Keshet and Schwarz (2014) or kind pronouns like *such* Schwager (2009))

- The distinction between intrapersonal and interpersonal anaphoric constructions in attitude reports also appears in the conventional wisdom readings of (1).

- *De re* reading-identifying scenarios require a singular interpersonal anaphoric specification and *de dicto* reading-identifying scenarios require an intrapersonal anaphor.

- Consider the textbook scenarios for *de re* (5) and *de dicto* (6) (von Fintel and Heim, 2011, adopted from p. 100, italics added).

(5)   I am walking along Newbury Street with Adrian. Adrian sees *a jacket* in a display window and wants to buy *it*. He tells me so. I don't reveal that I have one just like *it*. I report (1) to you.

(6)   Adrian's desire is to buy some jacket or other which fulfills the description that *it* is just like mine.

2.3   Intermediate Summary

- Fodor's reading shows that scope relationship reflects some of the aspects of attitude reports that anaphoric constructions can express, but not all.

- Thus, the role of inter- and intrapersonal anaphora in attitudes and their reports is the fundamental phenomenon to be explained by a semantic theory of the meaning of attitude reports.

- (Note: Fodor's puzzle is not the only case where anaphoric constructions create apparent scope paradoxes, compare e.g. Geach (1967)'s Hob-Nob sentences.)

3   The asymmetry of first- and third person perspective

- (1) ascribes a desire to Adrian from a third-person perspective whereas (3) is a report from Adrian's own first-person perspective.

- The first-person perspective grants introspection of one's own mental states – the attitudes themselves.

- A third person has no direct access to another person's mental states.

- If a justfication of a claim is a set of observations which, when taken together entail the truth of the claim, the asymmetry of first- and third person perspective finds expression in the way the truth of an attitude report is justified.

3.1   Justification of the truth of a first-person attitude report

- The truth of a first-person report like (3) is justified in virtue of the authority of the first person over what her state of mind is like.

- A rational justification (à la Davidson (1963)) for the self-ascription of a desire as in (3) would make explicit how the reported desire is parasitic on a set of beliefs: maybe Adrian saw a jacket like Malte's and liked it and thus wants to buy a jacket like the one he saw.

- The justification of a first-person report amounts to making explicit the intrapersonal anaphoric dependency of a parasitic attitude.

3.2   Justification of the truth of a third-person report

- Given that only the first person has authority over her state of mind, (1) cannot be justified to be true just in virtue of the reporter saying so.

- One strategy for a third person to justify the truth of (1) is when Adrian told the reporter (3)

- In this case, the third-person reporter can simply reproduce Adrian's self-ascription without taking on any further responsibility for why (1) is true besides the fact that Adrian said so.

- This justification of the truth of a third-person attitude report is called the *de dicto* strategy of justification: "the words which I, the speaker, am using to describe the attitude's content, are the same (at least as far as the relevant DP is concerned) as the words that the subject herself would use to express her attitude." (von Fintel and Heim, 2011, p. 84).

- Besides the delegation of responsibility to a first person *de dicto* (i.e. according to what the first person said), a third-person reporter can also justify (1) with an interpersonal anaphoric construction.

- The justification of the truth of an attitude report as being externally about the same object is called the *de re* justification: "[t]he term *de re* [...] indicates that there is a common object [...] whom I (the speaker) am talking about [...] and whom the attitude holder would be referring to if he were to express his attitude in his own words." (von Fintel and Heim, 2011, p. 84).

3.3   Intermediate Summary

- A semantic theory of the desires themselves and their reports should distinguish theorizing about the attitudes themselves from theorizing about attitude reports on the one hand and justification strategies for the truth of an attitude report (provided by reading-identifying scenarios) from the truth-conditions of an attitude report on the other

4   The attitudes themselves and reports of attitudes

- Heim (1992) popularized a semantic analysis of desire reports according to which "wanting something is preferring it to certain relevant alternatives, the relevant alternatives being those possibilities that the agent believes will be realized if he does not get what he want." (Stalnaker, 1984, p. 89).

- This quotation is not concerned with attitude reports but with the logical properties of attitudes from an agent's first-person perspective (e.g. the modelling of an agents beliefs about his beliefs or the ways in which an agent's beliefs may change in response to his experiences (Stalnaker, 1984, cp. p. 80)).

- Stalnaker discusses attitude ascriptions in a different section of the book.

- The basic problem resulting from the asymmetry of the first- and third-person perspective on attitudes with which (not only) Stalnaker is concerned is that to a third person, another agent's state of mind is basically a black box, the inputs and outputs to which can be observed but not its internal workings.

- A third-person attitude report like (1) must be based on observations about the behaviour of the reportee that the reporter considers to be indicative of his state of mind.

- Stalnaker is one of many philosophers who argued that an agent's behaviour indicates that she desires that P if "[t]o desire that P is to be disposed to act in ways that would tend to bring it about that P" (Stalnaker, 1984, p. 15).

- Acknowledging the asymmetry of first and third person, I propose to transfer Stalnaker'S conception of desires to the analysis of attitude reports as in (7)

(7)   A report of the form "x wants ϕ" is true if and only if x is disposed to take those actions which are likely to bring about a state of affairs of which the reporter is justified to believe that ϕ obtains.

- To the development of a formal account of desires as attitudes towards acting in a certain way, I take to be central what Stalnaker calls the forward-looking nature of desires, i.e. the fact that the reporter of (1) or (3) claims that the reportee acts towards a future state of affairs at which the descriptive content of the attitude report eventually becomes true.

- A formal framework in which such an analysis can be captured has been proposed in Singh and Asher (1993).

- The basic idea is to model time as a tree branching towards the future (in terms of CTL* Emerson (1990)) in which arc nodes are basic actions and maximal branches constitutes scenarios on which an agent acts

- (3) is true iff there is a set of scenarios which belong to Adrian's desire set such that if Adrian would act on any of these scenarios the reporter (i.e. Adrian) believes that the descriptive content of the report eventually becomes true.

- In (1), the reportee is not identical with the reporter and thus (1) is true iff there is a set of scenarios which belong to Adrian's desire set such that if Adrian would act on any of these scenarios the reporter (i.e. not Adrian) believes that teh descriptive content of the report eventually becomes true.

- (More details in the Appendix)

4.1   Intermediate Summary

- The truth-conditions of (1) do not reflect the conditions under which (1) is a true description of Adrian's attitudes themselves but the truth-conditions of (1) reflect the conditions under which (1) is a justified true description of Adrian's behaviour.

5   The justification of attitude reports

5.1   De dicto and de re justification

- In this section I consider in more detail the strategies which a reporter of (1) can employ to justify the truth of her belief that (1), where justified true belief amounts to an answer to the question put to the reporter "How do you *know* that (1)?"

- To do so, recall first the justification of first-person attitude reports.

- One strategy to justify the truth of (1) is to delegate the responsibility for the justification of the truth of (1) *de dicto* to Adrian himself uttering (3)

- Another way to justify the truth of (1) is with the help of an interpersonal anaphoric construction.

- A necessary condition for justifying (1) *de re* is that Adrian does not indicate his desire with a description as in (3) but with an interpersonal anaphoric description that indicates only the common object of interest while lacking descriptive content, recall the reading-identifying scenario for *de re* repeated in (8)

(8)   Adrian sees a jacket in a display window and wants to buy it. He *tells* me so.

- The *de re* justification of the truth of (1) amounts to resolving the singular anaphoric expression *it* occuring in Adrian's self-report "I want to buy it" to the antecedent description *a jacket like Malte's jacket*.

5.2   Fodor's puzzle

- The justification of the truth of (1) that puzzled Fodor is induced by complex interpersonal anaphoric constructions.

- The self-report of Adrian which underlies Fodor's reading involves a particular kind of complex anaphoric device of reference, e.g. as in *I want to buy one of those jackets* (cp. (4)).

- Compared to the interpersonal interpretation of the singular and simplex anaphoric expression *it*, what is special about *one of those* is that the expression involves not one but two anaphoric elements.

- On the one hand, Adrian's desire is directed towards an (unspecific) object that he (and only he himself) can identify with *one of*.

- On the other, there is a specific object that Adrian identifies with *those*.

- The problem that emerges from complex uses of anaphora is that they require the interpreter to distinguish the role of anaphora in self-attributions of attitudes themselves (the intrapersonal specification of *one of*) from the role of anaphora in attitude reports (the interpersonal specification of *those*).

- Consequently, it is not surprising that justifications that involve complex interpersonal anaphoric constructions are puzzling from the perspective of conventional wisdom.

- Modelling the meaning of attitude reports with the scope of *existential* quantifiers fails to acknowledge that interpersonal anaphoric justifications are not limited to singular anaphoric constructions but can also employ plural anaphoric constructions.

- Furthermore, it is because conventional wisdom conflates theorizing about the attitudes themselves with theorizing about their reports that it conceals the insight that anaphoric constructions do not only structure the attitudes themselves intrapersonally and attitude reports interpersonally, but that anaphoric constructions generally sit astride between attitudes and their reports in that they mediate between attitudes and their reports.

5.3   Intermediate Summary

- I take anaphoric constructions to be indicative of justification strategies themselves: no anaphoric construction – no justified attitude reporting.

5.4   Justification strategies

- If there is a truth at all to the claim that attitude ascriptions obtain validity from the very fact that we accept *de re* and *de dicto* justifications as being suitable licensors for our talk about attitudes, then justification strategies are part of the pragmatic conditions for the successful *use* of attitude ascriptions.

- Considering *de re* and *de dicto* justification as regulating our use of attitude ascriptions makes them part of the system of conventions according to which we ascribe attitudes

- It is a convention that a person has the descriptive authority over her own mental states and it is also a convention that a third-person report cannot claim authority over the description of other persons' mental states.

- (E.g., first-person attitude reports like (1) cannot be justified *de re* without giving up the social convention that a person has the authority over her own mental states.)

- The function of *de re* and *de dicto* justification in regulating the practice of attitude ascriptions is part of a theory of the pragmatics of attitude reports.

- Against this background, it should be explained how (1) is informative to an interpreter.

- One type of information an attitude report conveys is the rationalization of an agent's behaviour.

- But the more important communivative function of an attitude report is the way in which it serves as the input to the hearer's practical reasoning.

- The pro-attitudinal semantics for desire ascriptions that I presented is motivated on exactly these grounds: Singh and Asher (1993) argue in detail how the action-theoretic analysis of belief, intentions and desires provides the basis for the definition of a logic of rational agency in which agents can reason about themselves and other agents (and this is also the goal of Stalnaker's elaboration of inquiry).

6 Outlook

6.1 Extending the analysis to other non-doxastic attitudes

- The analysis of desire reports proposed can be modified so as to deal with other non-doxastic attitudes such as *hope* or *wish*

- The difference between *hope, wish* and *want* manifests itself in the strength with which they induce a tendency to act in the bearer of the attitude.

- While desires are understood as inducing a strong tendency to act in a certain way, a wish basically omits the forward-looking nature of desires and thus also can be used for goals which an agent cannot accomplish with his actions, like *I wish that I was never born.* vs. **I want that I was never born.*

- With respect to the relation of attitude and action, hopes seem to stand inbetween desires and wishes, as one the one hand, they share the forward-looking nature of desires but on the other also allow for outcomes of actions that an agent does not have under its own control like *I hope to pass the exams without preparation* vs. *I want to pass the exams without preparation.*

6.2 Extending the analysis to doxastic attitudes

- Like desires, an agent's beliefs can be observed from an external perspective only through the behaviour in which they manifest themselves under the hypothesis that "[t]o believe that P is to be disposed to act in ways that would tend to satisfy one's desires" (Stalnaker, 1984, p. 17).

- Like desires, beliefs are parasitic attitudes: "[b]eliefs have determinate content because of their presumed causal connections with the world" (Stalnaker, 1984, p. 17).

- It is the combination of both aspects of beliefs that should be considered to be fundamental to their reports, given that the justification of belief reports necessarily involves an anaphoric specification of how the reported belief is parasitic on its causal relation of acquaintance, compare e.g. (Quine, 1956, p. 179)'s famous Ortcutt scenario in (9) (italics added).

(9) "There is a certain man in a brown hat whom Ralph has *glimpsed* several times [...] Ralph suspects *he* is a spy. Also there is a gray-haired man, [...] whom Ralph is not aware of having *seen* except once at the beach. Now Ralph does not know it, but *the men* are one and the same."

6.3 Summary

- I argued that anaphoric constructions but not scope and intensional status of noun phrases are decisive for the meaning of attitude reports.

References

Davidson, D. (1963). Actions, reasons and causes. *Journal of Philosophy*, 60:695 – 700.

Emerson, E. A. (1990). Temporal and modal mogic. In van Leeuwen, J., editor, *Handbook of Theoretical Computer Science Vol. B*, Amsterdam. North-Holland Publishing Company.

Fodor, J. D. (1970). *The linguistic description of opaque contents.* PhD thesis, Massachusetts Institute of Technology, Dept. of Foreign Literatures and Linguistics.

Geach, P. T. (1967). Intentional identity. *Journal of Philosophy*, 64(20):627 – 632.

Heim, I. (1992). Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9:183–221.

Kamp, H. (1985). Context, thought and communication. *Proceedings of the Aristotelian Society*, 85:239–261.

Kamp, H., van Genabith, J., and Reyle, U. (2011). Discourse representation theory. In Gabbay, D. M. and Guenthner, F., editors, *Handbook of Philosophical Logic*, volume 15, pages 125 – 394. Springer, 2nd edition.

Keshet, E. and Schwarz, F. (2014). De re/de dicto (unpublished manuscript). Unpublished Manuscript.

Maier, E. (2015). Parasitic attitudes. *Linguistics and Philosophy*, 38(3):205–236.

Quine, W. V. O. (1956). Quantifiers and propositional attitudes. *Journal of Philosophy*, 53(5):177 – 187.

Romoli, J. and Sudo, Y. (2009). De Re / De Dicto Ambiguity and Presupposition Projection. In Riester, A. and Solstad, T., editors, *Proceedings of Sinn und Bedeutung 13*, pages 425 – 438.

Schwager, M. (2009). Speaking of qualities. In Cormany, E., Ito, S., and Lutz, D., editors, *Proceedings of SALT 19*.

Singh, M. P. and Asher, N. (1993). A logic of intentions and beliefs. *Journal of Philosophical Logic*, 22:513 – 544.

Stalnaker, R. C. (1984). *Inquiry*. MIT Press.

von Fintel, K. and Heim, I. (2011). Intensional semantics. MIT Spring 2011 Edition.

Appendix

- The basic idea spelled out in Singh and Asher (1993) is to include in a standard model theory of the language of Discourse Representation Structures (DRS) (see Kamp et al. (2011)) a time structure $\mathscr{T}$ and two functions $\mathbf{B}$ and $\mathbf{D}$. $\mathscr{T} = \langle \mathbf{T}, < \rangle$

- $\mathbf{T}$ is a set of possible times and $<$ a partial order on $\mathbf{T}$.

- We may view $\mathscr{T}$ as a tree-like structure of times branching towards the future in which arc labels are basic actions.

- We define a scenario S at a world w and time t as any maximal branch starting from t and $\mathbf{S}_{w,t}$ is the class of all scenarios at world w and time t.

- $\mathbf{B}$ and $\mathbf{D}$ are functions defined for a set of 'agents' (a subset of the set of individuals in the model theory) to subsets of $\mathbf{S}_{w,t}$ which assign beliefs $\mathbf{B}$ and desires $\mathbf{D}$ to agents at different worlds and times (qua modelling).

- (10) defines the interpretation of the DRS-predicate $DES(a,K)$, where the first argument is an agent and the second argument is a DRS K, as the set of scenarios S at t such that if the world w developed along any of them, the reporter b believes that the desire described with K would be realized at $t' > t$, i.e. that K has a verifying embedding g in M at w at t'.

(10)   Interpretation of $DES(a,K)$:
$[\![DES(a,K)]\!]_{M,w,t,g} = \{S | \exists w, t : S \in \mathbf{S}_{w,t} \wedge (\exists t', f : t' \in S \wedge g_w \subset_U f_w \wedge M \vDash_{w,t',f} BEL(b,K))\}$,
where a is the bearer of the desire and b the ascriber of the desire

(11)   Verifying embedding of $DES(a,K)$:
$M \vDash_{w,t,g} DES(a,K)$ iff $[\![DES(a,K)]\!]_{M,w,t,g} \in \mathbf{D}(w,t,a)$

Beliefs are defined relative to scenarios as the set of scenarios S at whose initial world w and time t, the descriptive content K of the belief is true under the given embedding g, cp. (12). I do not relativize belief ascriptions to the reporter-reportee distinction here, but see section 6 for additional discussion.

(12)   Interpretation of $BEL(a,K)$
$[\![BEL(a,K)]\!]_{M,w,t,g} = \{S | \exists w, t : S \in BEL(w,t,a) \wedge (\exists f : g_w \subset_U f_w \wedge M \vDash_{w,t,f} K))\}$

(13)   Verifying embedding of $BEL(a,K)$:
$M \vDash_{w,t,g} BEL(a,K)$ iff $[\![BEL(a,K)]\!]_{M,w,t,g} \in \mathbf{B}(w,t,a)$

- (14) is the semantic representation of the first-person report (3) in the representation language for attitudes proposed in Maier (2015).

- DRS_1 labels (for the sake of presentation) the representation for the global belief state of Adrian in which the desire representation labelled with DRS_2 is embedded.

- (14) is true iff there is a set of scenarios which belong to Adrian's desire set such that if Adrian would act on any of these scenarios the reporter (i.e. Adrian) believes that DRS_2 eventually becomes true.

(14)

- The representation of (1) in (15) involves a subtle but important difference.

- In (1), the reportee is not identical with the reporter and thus the semantics in (10) predicts that (1) is true iff there is a set of scenarios which belong to Adrian's desire set such that if Adrian would act on any of these scenarios the reporter (i.e. not Adrian) believes that DRS_3 eventually becomes true.

(15)